

crescent.cloud – Lessons Learned: From **Pipelines** to an **Interactive Web App**



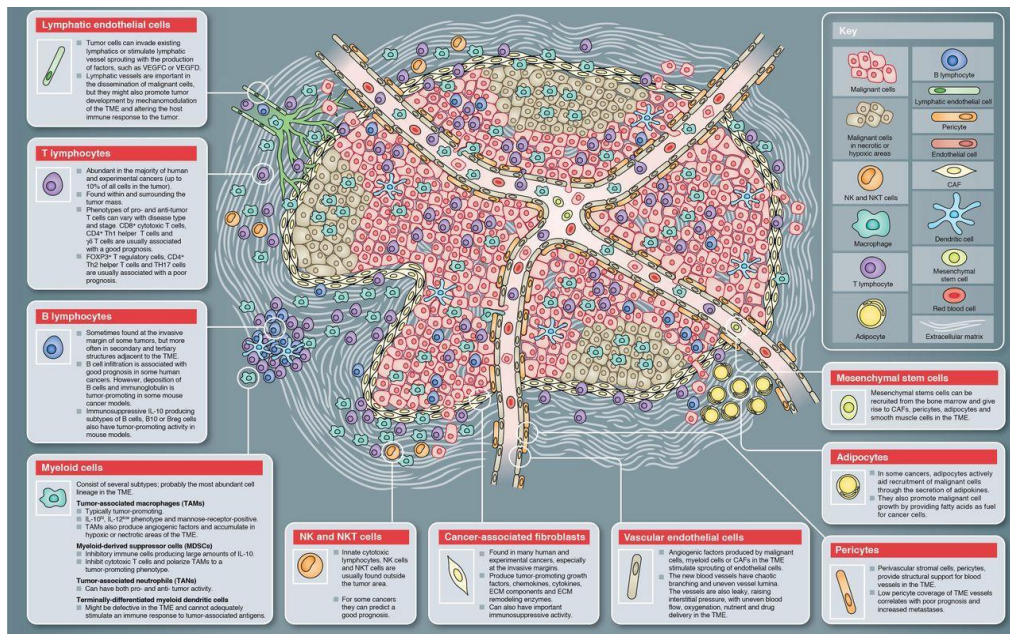
sułuxan mohanraj
martin do pham

Agenda

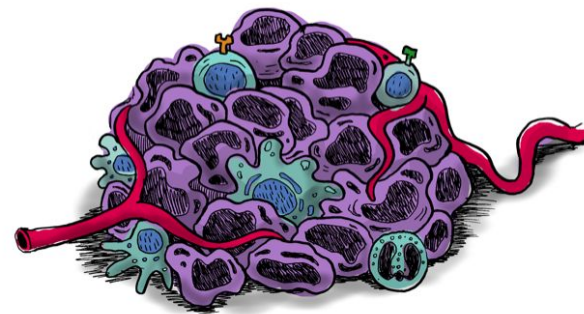
- **Motivation**
 - Tumour biology
 - Single-cell data
- **CReSCENT: CanceR Single Cell ExpressioN Toolkit**
 - Overview and current architecture
- **Lessons Learned**
 - Web app architecture
 - Workflow creation and execution
 - Prototyping for research
- **Next Steps**
 - Compute Canada cloud
 - Tool repository & pipeline builder

Motivation: Tumour Biology

Tumour Microenvironment



A Tumour...



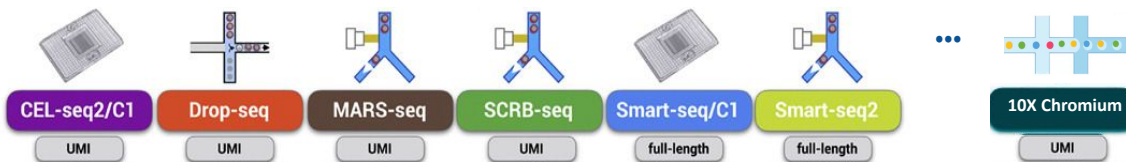
... More than just cancer cells!

(Balkwill, 2012)

(@pedromics)

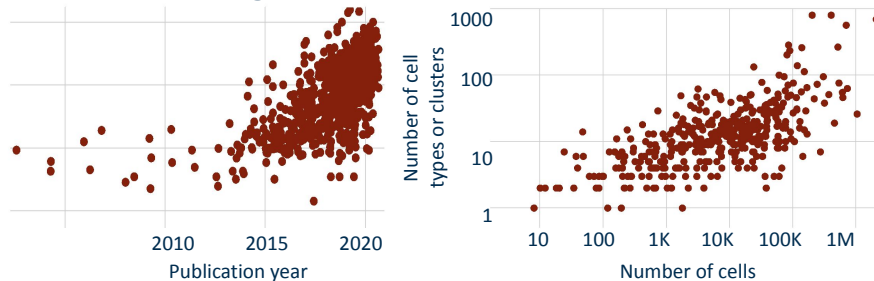
Motivation: Single-Cell Data

Single-Cell RNA-Seq Technologies



Modified from Ziegenhain (2017)

Single-Cell Data Growth

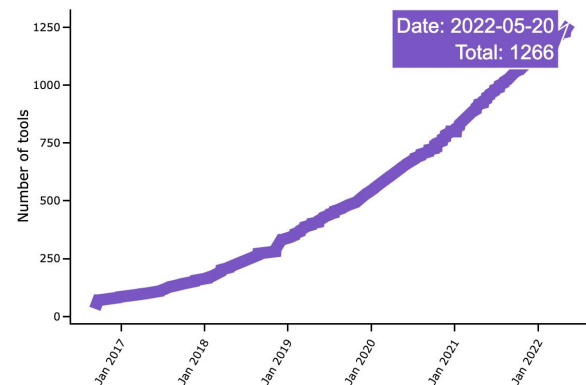
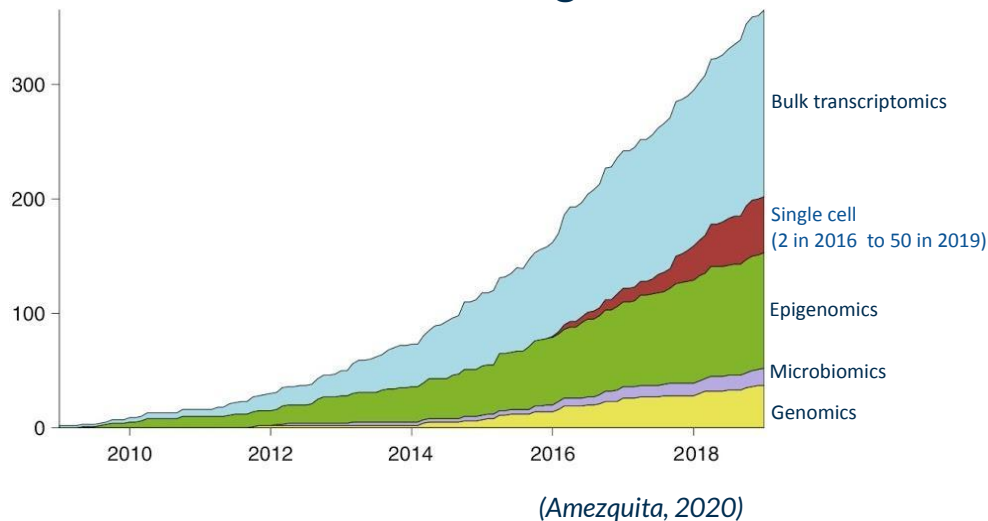


(Svensson, 2019)

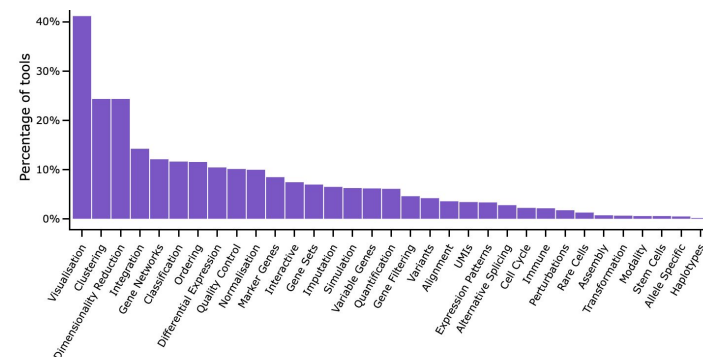
Motivation: Single-Cell Analysis

scRNA-tools.org

R/Bioconductor Packages



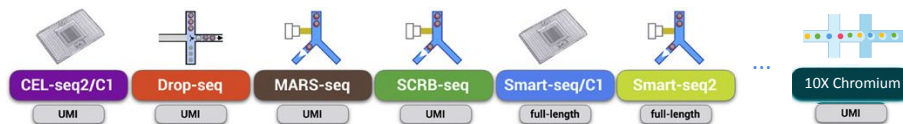
1266 tools...



...in over 30 categories

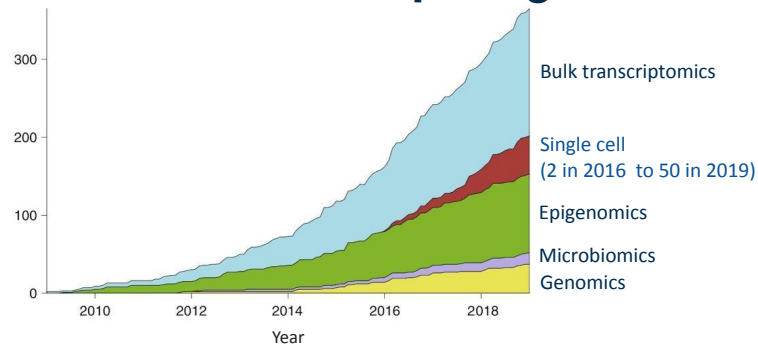
Motivation: Single-Cell Data

Single-Cell RNA-seq Technologies



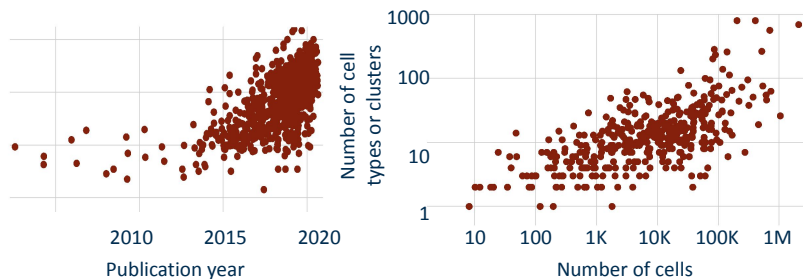
Modified from Ziegenhain (2017)

R/Bioconductor packages



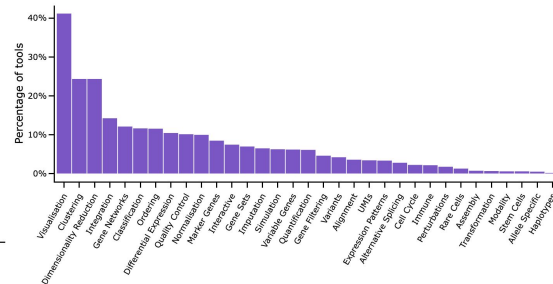
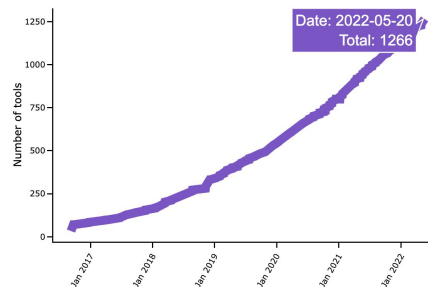
(Amezquita, 2020)

Single-Cell Data Growth



(Svensson, 2019)

scRNA-tools.org



1266 tools...

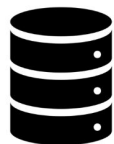
...in over 30 categories



single-cell
RNA-seq data



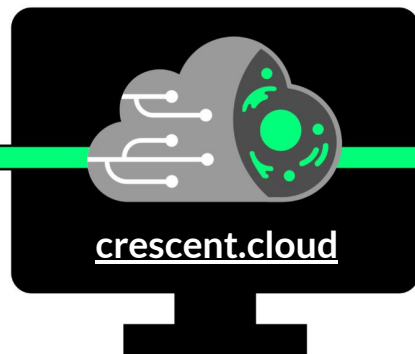
cell
metadata



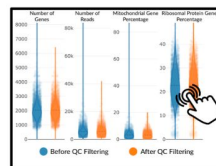
reference
datasets

CReSCENT

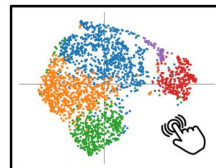
Cancer Single Cell Expression Toolkit



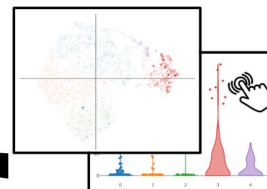
“single-cell analysis in the cloud”



quality
control



dimension
reduction



differential
gene expression



Nucleic Acids Research

CReSCENT: Cancer Single Cell Expression Toolkit

Suluxan Mohanraj, J. Javier Díaz-Mejía, Martin D Pham, Hillary Elrick, Mia Husić, Shaikh Rashid, Ping Luo, Prabhur Bal, Kevin Lu, Samarth Patel ... [Show more](#)
[Author Notes](#)

Nucleic Acids Research, Volume 48, Issue W1, 02 July 2020, Pages W372–W379,
<https://doi.org/10.1093/nar/gkaa437>

Published: 01 June 2020 [Article history](#) ▼

nature cancer

Gradient of Developmental and Injury Response transcriptional states defines functional vulnerabilities underpinning glioblastoma heterogeneity

Laura M. Richards, Owen K. N. Whitley, ... Trevor J. Pugh [✉](#) [+ Show authors](#)

Nature Cancer 2, 157–173 (2021) [Cite this article](#)

7637 Accesses | 18 Citations | 505 Altmetric | [Metrics](#)

Antitumor immune effects of preoperative sitravatinib and nivolumab in oral cavity cancer: SNOW window-of-opportunity study



Marc Oliva,^{1,2} Douglas Chepeha,³ Daniel V Araujo,⁴ J. Javier Díaz-Mejía,⁵ Peter Olson,⁶ Amy Prawira,⁷ Anna Spreafico,⁸ Scott V Bratman,^{8,9} Tina Shek,⁸ John de Almeida,³ Aaron R Hansen,³ Andrew Hope,^{8,9} David Goldstein,³ Ilan Weinreb,¹⁰ Stephen Smith,¹⁰ Bayardo Perez-Ordóñez,¹⁰ Jonathan Irish,³ Dax Torti,^{5,11} Jeffrey P. Bruce,³ Ben X. Wang,^{5,12} Anthony Fortuna,³ Trevor J. Pugh,^{5,11,13} Hrak Der-Torossian,^{1,4} Ronald Shazer,¹⁴ Nickolas Attanasio,¹⁵ Qingyan Au,¹⁶ Antony Tin,¹⁶ Jordan Feeney,¹⁶ Himanshu Sethi,¹⁶ Alexey Aleshin,¹⁶ Isan Chen,¹⁴ Lillian Siu,²

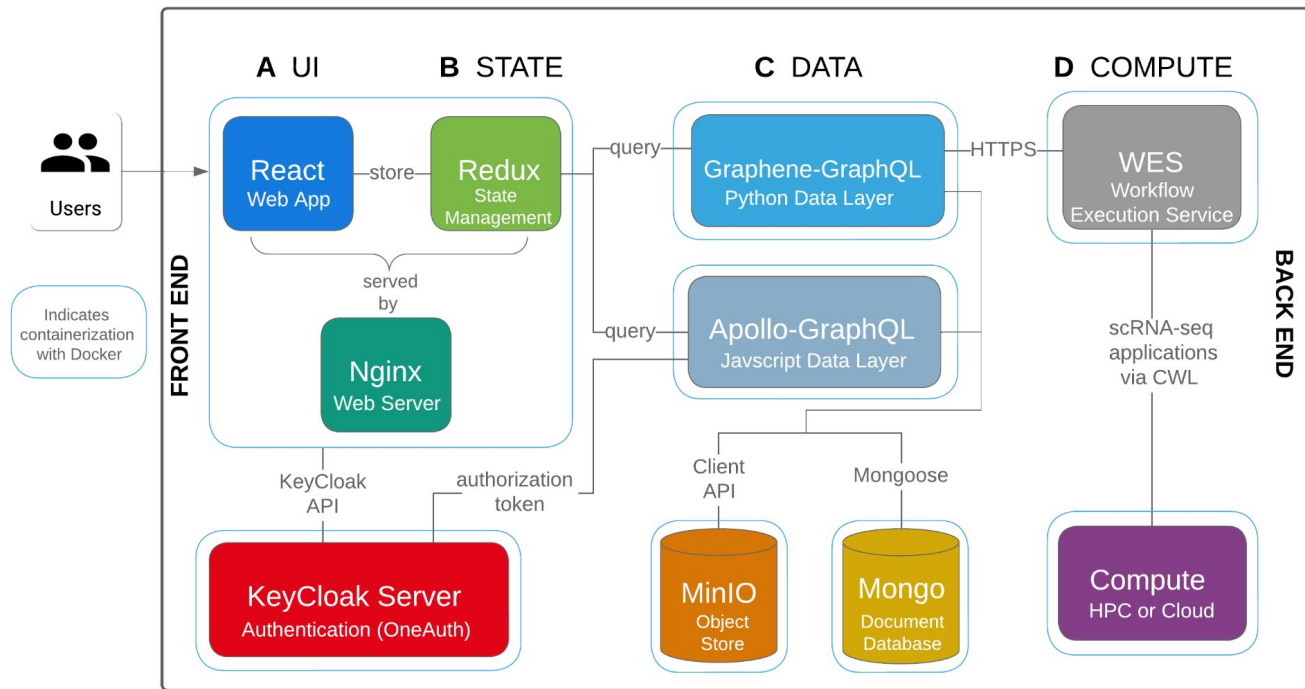
A comparison of data integration methods for single-cell RNA sequencing of cancer samples



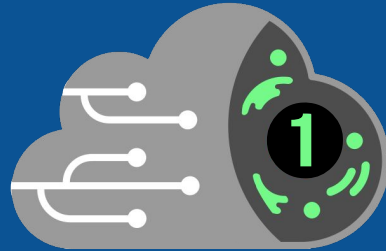
Laura M. Richards, Mazdak Riverin, Suluxan Mohanraj, Shamini Ayyadurthy, Danielle C. Croucher, J. Javier Díaz-Mejía, Fiona J. Coutinho, Peter B. Dirks, Trevor J. Pugh

doi: <https://doi.org/10.1101/2021.08.04.453579>

CReSCENT Architecture

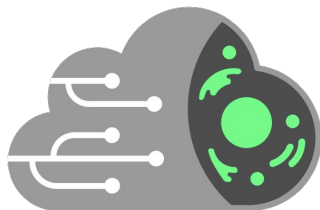


Lessons Learned: Web App Architecture



**keep a flexible
stack**

Flexible Stack leveraging GraphQL



M



E



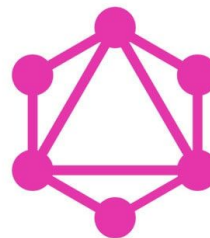
R



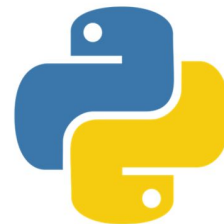
N



Apollo



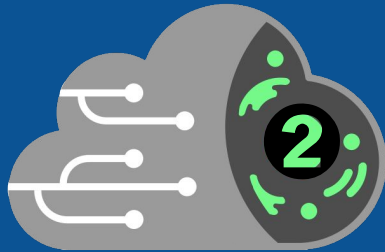
GraphQL



python™



Lessons Learned: Workflow Execution

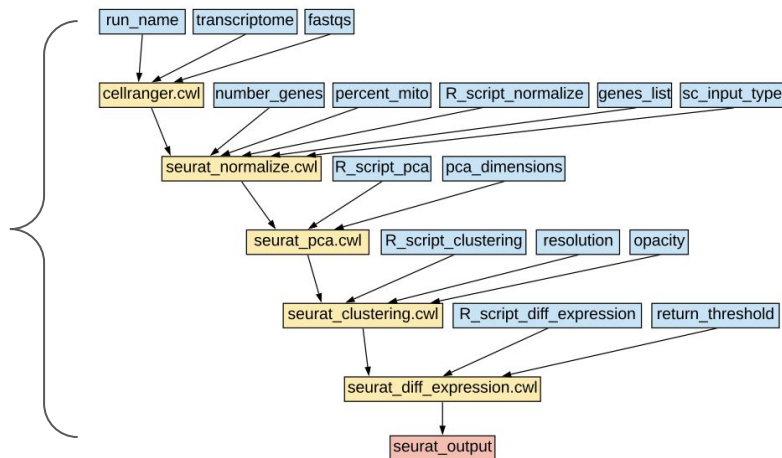


**use standardized
pipeline language and
workflow execution**

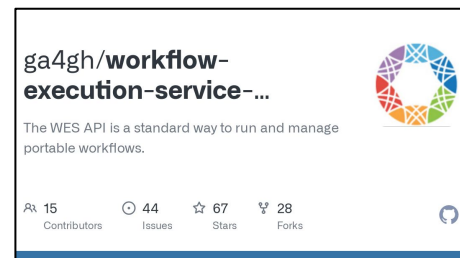
Standardized Pipelines and Workflow Execution



**Seurat single-cell
workflows
via R scripts**



**Common
Workflow
Language (CWL)**



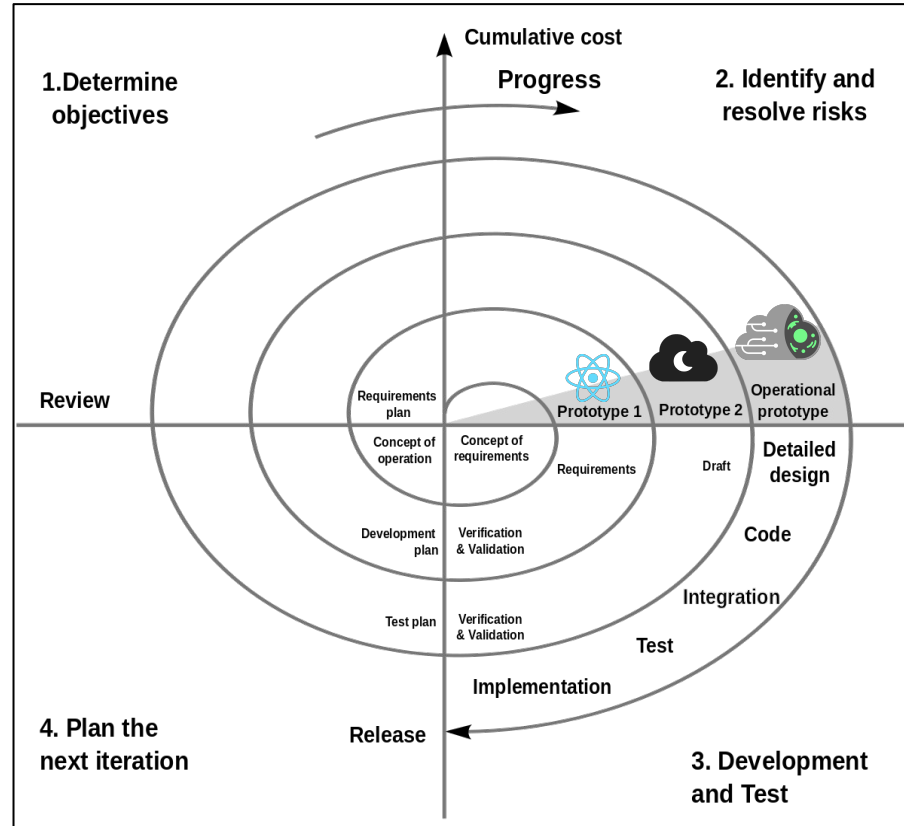
**Workflow
Execution Service
(WES API) from
GA4GH**

Lessons Learned: Prototyping for Research



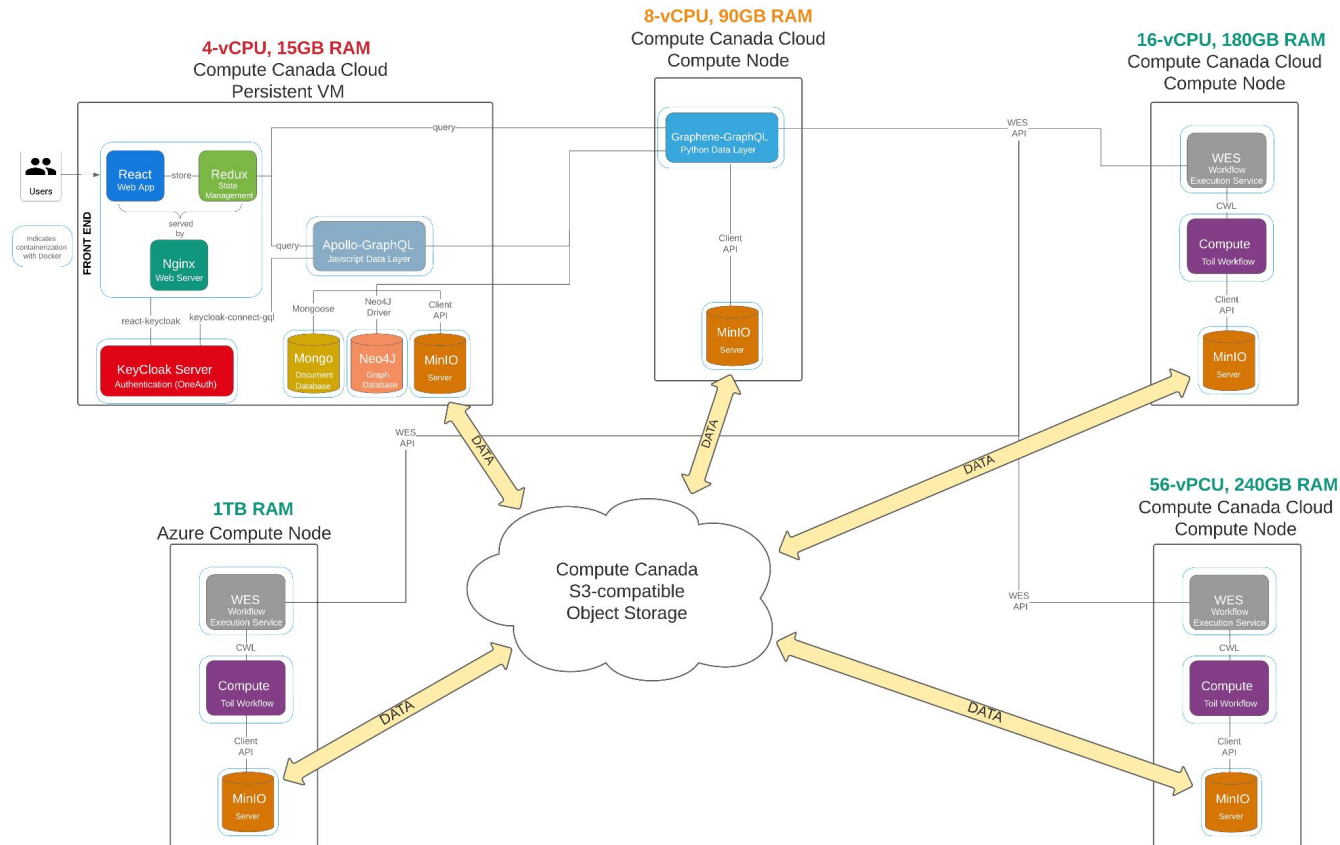
**spiral model of
development**

Prototyping for Research using the Spiral Model



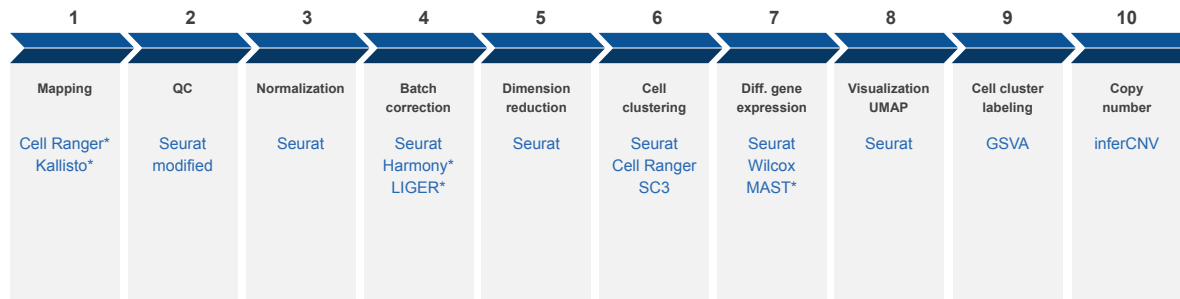
Modified from Boehm (1988)

Next Steps: Compute Canada Cloud

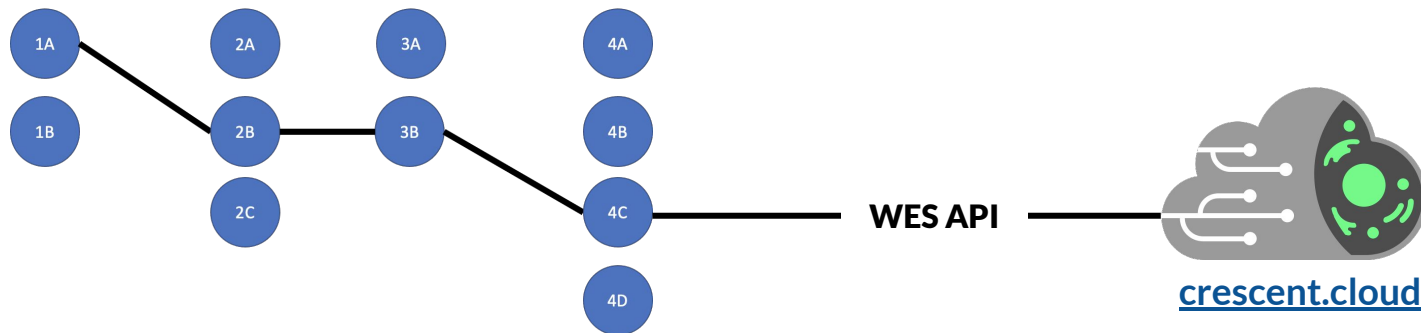


Next Steps: Tool Repository and Pipeline Builder

CReSCENT scRNA-seq Tools



Directed Acyclic Graph (DAG) CWL Pipeline Builder



Lessons Learned: Summary

1

keep a flexible
stack



2

use standardized
pipeline language and
workflow execution



3

spiral model of
development



Acknowledgements

CReSCENT Team

Pugh Lab (UHN)

Trevor Pugh

Suluxan Mohanrai

Javier Díaz-Mejía

Arnavaz Danesh

Ping Luo

Laura Richards

Samah El Ghamrasni

Danielle Croucher

Carlson Dultra

Brudno Lab (UHN DATA)

Michael Brudno

Martin Pham

Shaikh Rashid

Shooshtari Lab (UWO)

Parisa Shooshtari

Alaina Mahalanabis

Alaine Naidas

Erik Christensen

Centre for Computational Medicine (SickKids CCM)

Arun Ramani

Andrei Turinsky

Anna Guo

Grace Xu

Kevin Mo

Collaborators

PMGC (UHN)

Troy Ketela

Julie Tsao

Iulia Cirlan

Mandy Xu

Carl Virtanen

Zhibin Lu

Gary Bader (UofT)

Benjamin Haibe-Kains (UHN)

Mathieu Lupien (UHN)

Marc De Perrot (UHN)

Pam Ohashi (UHN)

Lillian Siu (UHN)

Ola Landgren (MSKCC)

Supported by



GenomeCanada



Ontario Genomics



CReSCENT Links

WEB APP: crescent.cloud

DOCS: pughlab.github.io/crescent-frontend

SCRIPTS: github.com/pughlab/crescent

CODEBASE: github.com/pughlab/crescent-frontend

NAR PAPER: doi.org/10.1093/nar/gkaa437

SLACK: sctoronto.slack.com

CONTACT US: crescent@uhnresearch.ca

